International Journal of Finance and Investment
ISSN: 3078-6002 (Print), ISSN: 3078-2295 (Online) | Vol. 4, No. 1, 2025

An Exploration of the Dual Impacts of Artificial Intelligence Credit Scoring on Banking System Risk

Yuxin Zhang

Monash business school (faculty of business and economics), Monash university, Melbourne, 3000, Australia

Abstract: The digital transformation of finance continues to deepen under the dual policy and practical drive of the implementation of the "Interim Measures for the Administration of Generative Artificial Intelligence Services" and the deepening promotion of inclusive finance. Artificial intelligence credit scoring has become the core technical support tool for commercial banks' risk management, and its impact on banking system risks presents a distinct dual nature. This article combines the practice of government-bank data collaboration with real-world commercial bank operations to systematically sort out the technical operating logic and risk action mechanism of artificial intelligence credit scoring: on the one hand, it analyzes its positive enabling value in mitigating traditional credit risks and expanding the coverage of inclusive financial services; on the other hand, it deeply analyzes new risk hazards induced by data collection bias, model "black box" characteristics, and cross-institutional systemic transmission of risks. The study found that after a joint-stock bank optimized its credit assessment model through multi-source heterogeneous data fusion, the non-performing loan loss rate dropped by 30 percentage points compared with before optimization; However, due to a lack of training data for rural customers and small and micro businesses, some banks experienced a misjudgment rate of up to 28% for normal transactions in rural areas. This article ultimately constructs a collaborative governance framework of "data governance—algorithm transparency—regulatory adaptation." This framework uses data governance to ensure source quality assessments, algorithm transparency to address model interpretation challenges, and regulatory adaptation to address technological iterations. This framework provides a practical path for balancing financial technology innovation and risk prevention and control, while also offering a reference for technology application by small and medium-sized banks and supporting the robust operation of the banking system.

Keywords: AI credit scoring, banking system risk, credit risk mitigation, model risk, regulatory technology.

1. Introduction

With the implementation of the Interim Measures for the Administration of Generative Artificial Intelligence Services and the deepening of digital cooperation between government and banks, AI credit scoring has transitioned from the technical exploration stage to the large-scale application stage. As of March 2025, more than 20 banks in my country have deployed large AI models for credit assessment. Changshu Rural Commercial Bank has built a new scoring system by integrating government affairs and financial data, providing 90,000 new citizens with pre-credit of 5.334 billion yuan, which directly reflects the value of technology empowerment. However, the Bank for International Settlements (BIS) warned that the complexity of AI algorithms may give rise to new forms of financial risks, and problems such as model black boxes and data bias cannot be ignored. Traditional credit scoring relies on limited financial data and has inherent shortcomings in the assessment of inclusive customer groups: new citizens and small and micro enterprises are often excluded due to lack of historical credit records, exacerbating the difficulty of financing. Artificial intelligence relies on machine learning to integrate multi-dimensional data. In theory, it can accurately portray risks, but in practice, the "herd effect" and algorithm discrimination that US SEC Chairman Gensler is worried about appear, increasing risk hazards [1]. Existing research focuses on the effects of technology application, lacks systematic discussion of the duality of risks, and lacks a feasible governance solution. Based on this, this article combines the real cases of Changshu Rural Commercial Bank and Industrial and Commercial Bank of China, and from the perspective of technical mechanisms, deeply analyzes the dual impact of artificial intelligence credit scoring on bank credit risk, operational risk and systemic risk, and proposes a governance path that is both innovative and secure, providing a practical reference for risk management in the digital transformation of banks.

2. Technical Mechanisms and Current Application Status of AI-powered Credit Scoring

Artificial intelligence credit scoring breaks through the limitations of traditional linear models. With the closed loop of “data fusion – feature mining – dynamic iteration”, it achieves iterative upgrades in risk assessment capabilities. Currently, its application has formed a diversified practice pattern [2]. At the technical architecture level, the system takes multi-source data as its core foundation, integrates financial data with external data such as government affairs, industry and commerce, and judicial affairs, and builds an assessment framework covering “basic attributes, social attributes, and social contributions”. Changshu Rural Commercial Bank and the local big data management bureau jointly established a laboratory to achieve multi-scenario reuse of scoring results through interface encapsulation. After being embedded in “Citizen Loan”, it not only completes the full online processing of credit, but also realizes the precise allocation of pre-credit. Its public data shows that the accuracy of pre-credit of this model has increased by 25%, and the bad debt rate is controlled below 1.2%. In terms of technical core, algorithms such as neural networks and random forests gradually replace traditional logistic regression models. Ant Group's Sesame Credit integrates user consumption trajectories and social behaviors to build a three-dimensional credit profile, enabling nearly 30 million users without traditional credit records to obtain first-time credit, improving the accuracy of risk characterization and service coverage. The application scenario presents the characteristics of "inclusive orientation + efficiency improvement". MyBank has served 40 million small and micro enterprises based on its "310" credit model. Traditional small and micro loan approvals take 3-5 working days, but this model reduces the time to seconds, increasing the approval rate by 40%, which directly reflects the value of AI in lowering the threshold. ICBC independently developed a large AI model to optimize risk control, and its non-performing loan ratio dropped from 1.73% to 1.42% in three years [3]. However, the industry faces a prominent problem of uneven development: large banks rely on their resource advantages to build complex models, while small and medium-sized banks, limited by data and computing power, often use third-party standardized models. These models are difficult to adapt to the localized credit needs of small and micro businesses in counties and rural customer groups, and can easily lead to regional credit "mismatches," posing a hidden danger for risk transmission. The lack of technical standards is a prominent bottleneck in the current industry. Different institutions have significant differences in model architecture design and verification process specifications. For example, some banks use hourly dynamic adjustment of feature weights, while others use a quarterly update model. Industry research shows that AI credit scores for companies with the same qualifications can vary by up to 30 points across different banks. This not only makes the assessment results lack comparability, but also directly affects the efficiency of cross-bank financing. The industry urgently needs to introduce unified standards to clarify the development direction.

3. The Mitigating Effect of AI-Powered Credit Scoring on Bank Risk

Artificial intelligence credit scoring, with its precise pricing, real-time warning and customer base expansion capabilities, effectively reduces the traditional risk exposure of the banking system and has become a key technology enabling means for risk management and control, and covers the entire process of credit approval and post-loan management. Its precise credit risk prevention and control effect is particularly outstanding [4]. Traditional credit scoring models rely too much on historical credit data and are difficult to effectively evaluate groups such as new citizens who have no credit records. AI technology can fill this assessment gap by mining the implicit credit value of non-financial data. Changshu Rural Commercial Bank has relied on this technology to grant 430 million yuan in credit to new citizens, and the actual scale of credit use has exceeded 170 million yuan. The non-performing loan rate of new citizens was controlled below 1.5%, achieving effective risk management while expanding the scope of service coverage. A joint-stock bank has specially deployed an AI real-time monitoring system to track corporate capital transactions, public opinion changes and other information 24 hours a day. Once risk signals such as delayed payment and executive changes are captured, the warning mechanism is immediately triggered, ultimately reducing the non-performing loan loss rate by 30%, fully demonstrating the practical application value of dynamic risk control. Operational risk and fraud risk have been significantly alleviated. The fraud detection system built based on machine learning technology can accurately identify abnormal transaction patterns and effectively detect false loan applications and feature collision attacks with the help of algorithms such as Isolation Forest [5]. The Context-Aware model developed by Goldman Sachs Group has increased the accuracy of financial report risk identification from 79% to 93% by retaining the rhetorical and logical features of financial report texts, reduced the omission of major risks by 68%, and shortened the financial report review cycle by 40%, reducing the subjective bias and efficiency loss in the manual review process. The value of risk reconstruction in the field of inclusive finance is very significant. AI technology allows banks to serve small and micro-enterprises that are excluded by traditional models. On the basis of precise risk control, Ant Financial has cumulatively issued loans of more than 10 trillion yuan and served more than 40 million small and micro-enterprises, successfully breaking the vicious cycle of "high risk - no courage to lend" in the industry. Federated learning technology further realizes cross-institutional data sharing under the premise of protecting data privacy. A regional bank alliance has increased the accuracy of small and micro-enterprise default prediction by 22% by jointly training AI models, and fully complies with the relevant requirements of the "Personal Information Protection Law" on data security.

4. New Risks in the Banking System Induced by AI-Powered Credit Scoring

In the process of improving risk assessment efficiency, artificial intelligence has also spawned new risks in the three dimensions of data, models, and decision-making, posing a potential threat to the stability of the banking system [6]. Data quality issues have made the risk of assessment bias more obvious. The accuracy of AI models is highly dependent on the quality of training data. The 2023 Federal Reserve Audit Report pointed out that if AI models are trained with outdated economic cycle data, the probability of corporate default will be overestimated by 37%-52% during an economic downturn. The problem of data collection bias is even more prominent. The anti-fraud system of a large bank uses urban customers as the main sample for training data, resulting in a 28% misjudgment rate for normal transactions in rural areas, which is 4.3 times that of urban areas. Semantic loss in unstructured data processing also affects assessment accuracy. JPMorgan Chase's NLP model has obvious shortcomings in identifying contingent liabilities in financial reports, which ultimately leads to an underestimation of environmental risk assessment results by 23%. The fragility of the model itself further aggravates operational and compliance risks. The "black box" nature of AI models makes the decision-making process difficult to explain. According to Visa's public statistics, 63% of customer complaints come from loan rejection decisions for which the reasons cannot be explained, and 29% of these complaints were verified to be caused by model misjudgment. The concealment of adversarial attacks further escalates the risk. The MITRE Red Team, by fine-tuning 27 non-critical data fields, caused a bank's risk control model's pass rate to increase by 15%, resulting in a direct loss of 2.3 million euros for a European bank in half a year. The model failure problem caused by concept drift is equally serious. During the 2022 cryptocurrency market crash, the traditional AI model's bad debt prediction error for related businesses was as high as 380%. Algorithm ethics issues will lead to reputational and legal risks for banks [7]. Algorithm discrimination has become explicit. The AI credit scoring system of a joint-stock bank shows that users who use domestic mobile phones priced below 1,000 yuan have a loan rejection rate 37% higher than other users. The US Consumer Financial Protection Bureau (CFPB) found that a bank's AI model had a 19% higher false rejection rate for African-American applicants than for white applicants. Such problems not only violate the principle of fair credit, but may also trigger regulatory penalties. A bank in the European Union was fined 8 million euros by the regulatory authorities due to the presence of proxy variable bias in its model, which sounded an ethical alarm for the entire industry [8].

In analyzing the new risks of AI credit scoring, this paper adopts the "multi-source evidence fusion + cross-scenario verification" methodology:

Data layer: Through regulatory documents such as the Federal Reserve audit report and EU penalty cases, combined with bank customer misjudgment rate surveys and enterprise model bias tests, the impact of data quality on assessment bias is quantified;

Model layer: Relying on Visa complaint attribution and MITRE red team adversarial experiments, the operational risks caused by the model's "black box" and vulnerability are verified;

Ethical layer: Comparing the US and European compliance frameworks (such as the CFPB investigation and the EU "Artificial Intelligence Act"), combined with micro cases (such as the relationship between mobile phone brands and credit decisions), the algorithm's ethical risk path is revealed.

5. Systemic Risk Transmission Paths of AI-Powered Credit Scoring

AI credit scoring, through the use of model homogeneity and risk transmission across institutions, amplifies local risks into systemic threats, further increasing the vulnerability of the banking system. Model homogeneity can easily lead to the risk of "herd effect". Currently, when building AI credit scoring models, most banking institutions use similar data sources and algorithm frameworks. Due to the lack of independent research and development capabilities, some small and medium-sized banks directly purchase third-party standardized models, forming the "single culture" problem warned by US SEC Chairman Gensler - that is, the risk judgment logic of the entire industry is highly similar. If the core model makes a misjudgment, it may cause multiple banks to simultaneously shrink credit or centralize credit: In 2023, the US banking system generated 12 million unnecessary transaction verifications because the AI model did not adjust the off-site login weight parameters in a timely manner. This not only increased customer operating costs, but also directly affected the short-term market liquidity supply. This kind of coordinated misjudgment is more likely to cause systemic shocks in the banking system during an economic downturn. Cross-institutional data sharing will aggravate risk contagion [9]. Although government-bank cooperation and industry data alliances can improve the accuracy of model evaluation through data complementarity, they also allow risks to break through the firewall boundaries of a single institution. A regional banking alliance was affected by the mixing of abnormal transaction records in shared data. These records originated from historical invalid merchant data that had not been cleaned up by a member bank and were directly circulated due to the lack of a pre-data verification link in the sharing mechanism. This caused the risk control models of five member banks to make misjudgments at the same time, and the number of new loan rejections in a single day increased by three times compared to normal days, which directly confirmed the chain effect of risk transmission across institutions. Although data sharing technologies such as blockchain can ensure the credibility of the data transmission process, they cannot prevent the spread of original data deviations among multiple institutions, which further amplifies the scope and impact of risk transmission. From the perspective of technical implementation, the "blockchain + proxy re-encryption" solution proposed by Su Zhe et al. (2020) uses distributed ledgers to ensure the integrity of data transmission and achieves privacy protection through proxy re-encryption technology, but it still does not cover the original data verification link, so it cannot solve the problem of cross-institutional spread of data deviations [10]. Furthermore, the risk of ambiguous ownership in data sharing can easily lead to hidden disputes. In one provincial government-bank data collaboration project, the bank's model trained on this data was found to be infringing due to a lack of clear commercial authorization for government data. The bank was not only required to remove the relevant scoring system from the platform but also to bear 2.3 million yuan in compensation. This risk further increases the compliance costs and uncertainty of cross-institutional collaboration.

Lagged supervision can further amplify systemic risks. my country's current "Regulations on the Administration of the Credit Reporting Industry," drafted before the widespread adoption of AI, does not address the division of responsibilities for AI credit scoring models. The "Interim Measures for the Administration of Generative Artificial Intelligence Services" focus on service compliance and have limited regulatory effectiveness for model risks in the financial sector. This leaves the central bank facing a dilemma when conducting AI financial regulation. In contrast, the EU's Artificial Intelligence Act includes credit assessment in the scope of high-risk AI applications, and builds a full life cycle supervision system from data collection, model training to result application; while the progress of my country's regulatory technology construction lags behind the speed of technology application. Although some local regulatory authorities have piloted AI risk monitoring tools, the lack of a national unified data interface standard makes it difficult for monitoring data to be interoperable. It is difficult to achieve real-time and comprehensive monitoring of AI model risks, which significantly increases the difficulty of preventing and controlling systemic risks [11].

6. Risk Prevention and Control Pathways and Policy Recommendations Under Dual Impacts

Balancing the innovative value of AI-powered credit scoring with its potential risks requires building a comprehensive governance system encompassing "data foundation, algorithm transparency, and regulatory oversight" to synergize technological empowerment with risk prevention and control. Strengthening the foundation of data governance is paramount in risk prevention and control. Cross-departmental data sharing mechanisms should be implemented to integrate core data resources such as finance, taxation, and social security. Data collection standards and specifications should also be clarified to avoid assessment bias caused by incomplete sample coverage. The Changshu Rural Commercial Bank's model of government-bank data collaboration offers valuable insights. By connecting to local government data platforms to obtain authoritative data, it ensures data authenticity and reduces human bias in the collection process. Furthermore, data desensitization and differential privacy technologies need to be promoted. HSBC uses the Bayesian CPD method to encrypt sensitive customer information while balancing data security and model evaluation accuracy. Promoting algorithm transparency and explainability is key to resolving these issues. Financial institutions should deploy Explainable AI (XAI) tools. For example, FICO's Explainable AI Suite uses decision trajectory diagrams to visually display the influence of each feature on the scoring result, reducing customer dispute resolution time by 72%. It is recommended to establish an algorithm filing system requiring banks to disclose key information such as model architecture and feature weightings. The "credit-behavior" combination model piloted in the Yangtze River Delta region can serve as an industry reference model. At the same time, a fairness verification mechanism should be embedded in model training. Citibank, through the FairML framework, reduced the ROC difference between applicants of different ethnicities from 0.15 to 0.03, achieving a balance between efficiency and fairness. Building an adaptive regulatory system and industry self-discipline mechanism is an important guarantee [12]. Regulatory authorities need to speed up special legislation in the field of AI finance, clarify the division of model responsibilities and risk red lines, and promote the "regulatory sandbox" pilot, such as conducting inclusive testing on the dynamic weight adjustment model of Ant Group. It is also necessary to simultaneously develop regulatory technology tools to achieve real-time risk monitoring of AI models. The algorithm risk assessment framework of the Bank for International Settlements can provide technical support. In addition, the role of industry associations should be fully utilized to formulate unified model evaluation standards and verification processes to solve the industry problem of lack of comparability of evaluation results of different institutions.

7. Conclusion

The dual impact of AI credit scoring on the risk of the banking system is essentially a dialectical unity of technological innovation and risk prevention and control: it not only provides accurate and efficient tools for bank risk management, but also generates new risks at the data, model and system levels. Practice has fully proved that this technology has helped institutions such as the Industrial and Commercial Bank of China to achieve a steady decline in non-performing loan rates through multi-source data fusion and dynamic risk monitoring. Changshu Rural Commercial Bank has also used this to break through the financing bottleneck of new citizens and play a core role in mitigating credit risks and promoting the development of inclusive finance. However, at the same time, the difference in the misjudgment rate of urban and rural transactions caused by data collection bias, the compliance disputes caused by the "black box" model, and the "herd effect" caused by model homogeneity [13] have become potential hidden dangers threatening the stability of the banking system. In 2023, the US banking system caused 12 million unnecessary transaction verifications due to the delayed adjustment of the model's cross-regional login weights, and the 8 million euro fine issued by the European Union for bank algorithm discrimination, all of which directly confirm the reality of these risks. Balancing these dual impacts requires collaborative efforts from multiple parties: banking institutions should improve data governance mechanisms, promote algorithm transparency, and integrate fairness and explainability throughout the model lifecycle; regulators should accelerate legislation specifically for AI finance and simultaneously advance the development of regulatory technology to establish a "technology-adaptive" regulatory system; and the industry should establish unified model evaluation standards and self-regulatory norms. Only by fostering a synergistic approach of "banking practice, regulatory guidance, and industry collaboration" can the technical value of AI credit scoring be fully unleashed, systemic risks effectively mitigated, and the high-quality development of the banking system be firmly supported. Future research could further focus on the current predicament faced by small and medium-sized banks. These institutions, constrained by data reserves and computing power, face even greater challenges in adapting AI credit scoring models. Therefore, it is necessary to explore low-cost, easily implemented risk prevention and control approaches, ultimately promoting the deep integration of inclusive finance and financial stability.

References

[1] Kash D, Tkachenko N. Risks of Artificial Intelligence in Credit Rating [J]. Springer Books, 2025.

[2] Chen Yijun, Gao Haoran, Ding Zhijun. Credit Assessment Model Based on Dynamic Machine Learning [J]. Computer Science, 2023, 50(01): 59-68.

[3] Li Xiaoqing. Construction and Application of Bank Big Data Intelligent Risk Control Platform [J]. China Financial Computer, 2019, (04): 71-74. DOI:CNKI:SUN:ZGJN.0.2019-04-017.

[4] Faheem M A. Artificial Intelligence-driven Risk Assessment Model: Revolutionizing Credit Scoring and Default Prediction [J]. Landmark Research and Engineering Journal, 2021, 5(3): 177-186.

[5] Lü Fang, Tang Fenghe, Huang Junheng, et al. Research on Financial Fraud Account Detection for Imbalanced Datasets [J]. Computer Engineering, 2021, 47(06): 312-320. DOI: 10.19678/j.issn.1000-3428.0058006.

[6] Locatelli R, Pepe G, Salis F. Artificial Intelligence and Credit Risk [J]. Springer Books, 2022.

[7] Cash D, Tkachenko N. Risks of Artificial Intelligence in Credit Rating [J]. Springer Books, 2025.

[8] +4 Kleinberg J, Mullainathan S, Raghavan M. The Inherent Trade-offs in Fair Risk Scoring [J]. arXiv preprint arXiv: 1609.05807, 2016.

[9] Mao Xuncheng, Wang Zhouwei. Research on the Network Contagion Effect of Systemic Credit Risk [J]. Journal of Beijing Institute of Technology (Social Sciences Edition), 2014, 16(04):57-63+130. DOI:10.15918/j.jbitss1009-3370.2014.04.021.

[10] Su Zhe, Wang H, Wang H, et al. A Financial Data Security Sharing Solution Based on Blockchain Technology and Proxy Re-encryption Technology [C]//2020 IEEE Third International Conference on Industrial Safety and Informatization (IICSPI). IEEE, 2020: 462-465.

[11] Kumar G, Rahman MR, Rajverma A et al. Predicting bank systemic risk: A machine learning approach [J]. Journal of Management Modeling, 2024, 19(2): 441-469.

[12] Zeng Fanrong. Application and impact of artificial intelligence in the financial industry [J]. International Finance, 2018, (04): 36-42. DOI:10.16474/j.cnki.1673-8489.2018.04.007.

[13] Barocas S, Selbst A D. Big data's disparate impact [J]. Calif. L. Rev., 2016, 104: 671.